Mon. Not. R. Astron. Soc. **399,** L128–L131 (2009) doi:10.1111/j.1745-3933.2009.00738.x

The scale of homogeneity of the galaxy distribution in SDSS DR6

Prakash Sarkar,[1⋆] Jaswant Yadav,[2⋆] Biswajit Pandey[3⋆] and Somnath Bharadwaj[1⋆]

[1]*Department of Physics & Meteorology and Centre for Theoretical Studies, Indian Institute of Technology, Kharagpur 721302, India*
[2]*Harish Chandra Research Institute, Chhatnag Road, Jhunsi, Allahabad 211019, India*
[3]*Inter-University Centre for Astronomy and Astrophysics, Pune 411007, India*

Accepted 2009 August 7. Received 2009 July 20; in original form 2009 June 5

ABSTRACT

The assumption that the Universe, on sufficiently large scales, is homogeneous and isotropic is crucial to our current understanding of cosmology. In this Letter, we test if the observed galaxy distribution is actually homogeneous on large scales. We have carried out a multifractal analysis of the galaxy distribution in a volume-limited subsample from the Sloan Digital Sky Survey (SDSS) Data Release 6. This considers the scaling properties of different moments of galaxy number counts in spheres of varying radius, r, centred on galaxies. This analysis gives the spectrum of generalized dimension $D_q(r)$, where $q > 0$ quantifies the scaling properties in overdense regions and $q < 0$ in underdense regions. We expect $D_q(r) = 3$ for a homogeneous, random point distribution. In our analysis, we have determined $D_q(r)$ in the range $-4 \leq q \leq 4$ and $7 \leq r \leq 98\,h^{-1}$ Mpc. In addition to the SDSS data, we have analysed several random samples which are homogeneous by construction. Simulated galaxy samples generated from dark matter N-body simulations and the Millennium Run were also analysed. The SDSS data is considered to be homogeneous if the measured D_q is consistent with that of the random samples. We find that the galaxy distribution becomes homogeneous at a length-scale between 60 and 70 h^{-1} Mpc. The galaxy distribution, we find, is homogeneous at length-scales greater than 70 h^{-1} Mpc. This is consistent with earlier works which find the transition to homogeneity at around 70 h^{-1} Mpc.

Key words: methods: numerical – galaxies: statistics – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

The assumption that the universe is homogeneous and isotropic, known as the cosmological principle, is perhaps the most fundamental postulate of the currently well accepted standard model of cosmology (e.g. Weinberg 2008). The Sloan Digital Sky Survey (SDSS), the largest galaxy survey till date, has mapped out the galaxy distribution in a large volume of space. This provides a unique opportunity to test the cosmological principle observationally. The two-point correlation function and the power spectrum, which are the standard techniques to quantify galaxy clustering (Peebles 1980), rely on the assumption that the cosmological principle is valid on a sufficiently large scale, and hence cannot be used to test the cosmological principle.

Coleman & Pietronero (1992) have proposed that the universe has a fractal structure. If true, this violates the cosmological principle. While some of the subsequent analysis of galaxy surveys show evidence for homogeneity in the galaxy distribution at sufficiently large scales (Guzzo 1997; Bharadwaj, Gupta & Seshadri 1999; Pan & Coles 2000; Tikhonov, Makarov & Kopylov 2000; Kurokawa, Morikawa & Mouri 2001), there are others who either fail to find a transition to homogeneity or find definite evidence for a fractal structure out to the largest scale probed (Amendola & Palladino 1999; Hatton 1999; Best 2000; Baryshev & Bukhmastova 2004).

In a recent study, Yadav et al. (2005) have analysed nearly two-dimensional (2D) strips from the SDSS Data Release 1 to find a definite transition to homogeneity occurring at the length-scale 60–70 h^{-1} Mpc. Hogg et al. (2005) have analysed the distribution of luminous red galaxies (LRG) in the SDSS to find a transition to homogeneity at $\sim 70\,h^{-1} Mpc$. The transition to homogeneity, however, is contested by Sylos Labini, Vasilyev & Baryshev (2007) and Sylos Labini et al. (2009a, b) who fail to find a transition to homogeneity at the largest scales ($\sim 100\,h^{-1}$ Mpc) which they probe in the 'Two-Degree Field (2dF)' and SDSS galaxy surveys.

In this Letter, we have analysed the SDSS Data Release 6 (DR6) to test if the galaxy distribution exhibits a transition to homogeneity at large scales, and if so to determine the scale of homogeneity. To achieve this, we have carried out a multifractal analysis (Martinez & Jones 1990; Borgani 1995) of the galaxy distribution in a volume-limited subsample drawn from the SDSS DR6. The

⋆E-mail: prakash@cts.iitkgp.ernet.in (PS); jaswant@hri.res.in (JY); biswa@iucaa.ernet.in (BP); somnathb@iitkgp.ac.in (SB)

multifractal analysis quantifies the scaling properties of a number of different moments of the galaxy counts, allowing for the possibility that this may differ depending on the density environment. The current work improves upon the earlier 2D analysis of Yadav et al. (2005). The large contiguous region covered by SDSS DR6 allows us to carry out a full three-dimensional analysis covering scales up to 103 h^{-1} Mpc in the volume-limited sample. The transition to homogeneity has been identified by comparing the observations with a random point distribution having the same number and geometry as the actual galaxy distribution. The observations have also been compared with Λ cold dark matter (ΛCDM) N-body simulations which are also used to determine the error bars.

A brief outline of the Letter is as follows. We describe the data and the method of analysis in Section 2. Section 3 presents the results and conclusions of our study.

2 DATA AND METHOD OF ANALYSIS

2.1 SDSS DR6 data

Our present analysis is based on galaxy redshift data from the SDSS DR6 (Adelman-McCarthy et al. 2008). The SDSS DR6 includes 9583 deg^2 of imaging and 7425 deg^2 of spectroscopy with 790 860 galaxy redshifts. For the present work, we have used the main galaxy sample for which the target selection algorithm is detailed in Strauss et al. (2002). The main galaxy sample comprises of galaxies brighter than a limiting r-band Petrosian magnitude 17.77. The data were downloaded from the Catalog Archive Server (CAS) of SDSS DR6 using a Structured Query Language (SQL) search. We have identified a contiguous region in the Northern Galactic Cap which spans $-50° < \lambda < 30°$ and $-6° < \eta < 35°$, where λ and η are survey co-ordinates defined in Stoughton et al. (2002). A volume-limited galaxy subsample was constructed in this region by restricting the extinction corrected Petrosian r-band apparent magnitude to the range $14.5 \leq m_r \leq 17.77$ and restricting the absolute magnitude to the range $-20 \geq M_r \geq -21$. This gives us 41 234 galaxies in the redshift range $0.04 \leq z \leq 0.11$ which corresponds to comoving radial distances from 130 to 335 h^{-1} Mpc.

2.2 *N*-body data

We have simulated the dark matter distribution at $z = 0$ using a particle-mesh (PM) N-body code. This was used to simulate the observed galaxy distribution by identifying randomly chosen dark matter particles as galaxies. Here, we have chosen 41 234 particles from a region which is exactly identical to the survey volume.

The N-body simulations use 512^3 particles on a 512^3 mesh, with grid spacing 1 h^{-1} Mpc. The simulations use a ΛCDM power spectrum with cosmological parameters $(\Omega_{m0}, \Omega_{\Lambda 0}, h, n_s, \sigma_8) = (0.274, 0.726, 0.705, 0.96, 0.812)$ (Komatsu et al. 2009). The effect of peculiar velocities was incorporated using the plane parallel approximation. We have run five independent realizations of the N-body code, and the simulated galaxy distributions were analysed in exactly the same way as the actual data.

The galaxies in our simulations are assumed to trace the dark matter distribution exactly. Galaxy formation is a complex non-linear process, and it is quite possible that the galaxies are a biased tracer of the underlying dark matter distribution. To account for this possibility, we have also considered a semi-analytic galaxy catalogue (Croton et al. 2006) from Millennium Run simulation (Springel et al. 2005). Semi-analytic models are simplified simulations of the formation and evolution of galaxies in a hierarchical clustering scenario incorporating all relevant physics of galaxy formation processes. The spectra and magnitude of the model galaxies were computed using population synthesis models of Bruzual & Charlot (2003), and we use the catalogue where the galaxy magnitudes are available in the SDSS u, g, r, i, z filters. The catalogue contains about nine million galaxies in a $(500\,h^{-1}\,\mathrm{Mpc})^3$ box. Using the peculiar velocities, we map the galaxies to redshift space and then identify a region having the same geometry and choose the same number of galaxies within the specified magnitude range as the actual galaxy sample.

2.3 Method of analysis

Scale invariance of galaxy clustering (Jones et al. 2005), at least at small scales, motivates us to use fractal analysis. There are various definitions of fractal dimension, for example, the Hausdorff dimension, capacity dimension, similarity dimension, box counting dimension, etc. These definitions represent particular cases of the multifractal spectrum of generalized dimensions. In this Letter, we use the generalized Minkowski–Bouligand dimension to characterize the galaxy distribution. This contains complete information about various moments of the galaxy counts, and is, therefore, well suited to multifractal analysis.

Considering the i th galaxy as centre, we determine $n_i(<r)$ the number of other galaxies within a sphere of comoving radius r. The generalized correlation integral is defined as

$$C_q(r) = \frac{1}{MN} \sum_{i=1}^{M} [n_i(<r)]^{q-1}, \tag{1}$$

where M is the number of centres, N is the total number of galaxies and $q - 1$ refers to a particular moment of the galaxy counts. For a fixed r, we have used all the galaxies, barring those that lie within a distance r from the survey boundary, as centres. In our analysis, $N = 41234$ and $M = 27706$ and 6108 for $r = 20$ and 70 h^{-1} Mpc, respectively.

The generalized Minkowski–Bouligand dimension D_q follows from the correlation integral, and is given by

$$D_q(r) = \frac{1}{q-1} \frac{\mathrm{d} \log C_q(r)}{\mathrm{d} \log r}. \tag{2}$$

For $q > 1$, D_q probes the scaling behaviour of galaxies in high-density environments, for example clusters and super-clusters, and for $q < 1$, D_q probes the same in underdense environments like voids. We have varied q from $q = -4$ to $q = +4$ in steps of 1. In principle, we could have considered even higher and lower values of q but the finiteness of the data increases the scatter in the D_q value as the value of $|q|$ increases (Bagla, Yadav & Seshadri 2008). We have calculated $D_q(r)$ by numerically differentiating C_q using Ridders' method (Press et al. 1992) considering $C_q(r)$ at three consecutive r values, each separated by 5 h^{-1} Mpc. Using a smaller r interval (1 h^{-1} Mpc) , we find an increase in the fluctuations in $D_q(r)$ while using a larger interval (10 h^{-1} Mpc) gives nearly the same results as (5 h^{-1} Mpc).

We have determined $C_q(r)$ in the range $1 \leq r \leq 103\,h^{-1}$ Mpc and $D_q(r)$ in the range $7 \leq r \leq 98\,h^{-1}$ Mpc for the data as well as the simulations. To assess the transition to homogeneity, we have also generated 'random samples' which contain randomly located points. The number of points and the volume coverage of these samples is exactly same as that of SDSS data. The point distribution in the random samples are homogeneous by construction, and we have generated 10 independent samples. The data, or the simulations, are considered to be homogeneous on length-scales where $D_q(r)$ is consistent with that of the random sample.

3 RESULTS AND CONCLUSIONS

Fig. 1 shows the correlation integral $C_q(r)$ for $q = -2$ and $+3$. The behaviour is similar for other values of q. We find that $C_q(r)$ increases with r for positive values of q whereas it falls progressively for negative values of q. It is quite clear that at all length-scales the SDSS galaxies, the dark matter and the Millennium simulations all exhibit the same scaling behaviour. It is also quite clear that on small scales, less than approximately 30 and 40 h^{-1} Mpc for $q = -2$ and 3, respectively, this scaling behaviour is distinctly different from that of the random samples. The scaling behaviour of the actual data, simulations and the random sample all appear to match as $r \rightarrow 100\,h^{-1}$ Mpc, and the $C_q(r)$ curves all coincide if plotted on the same scale. The match between the data and the random samples indicates a transition to homogeneity at a length-scale smaller than 100 h^{-1} Mpc, and the issue is to identify the length-scale where this transition occurs.

As discussed earlier, we have numerically differentiated $C_q(r)$ to calculate the Minkowski–Bouligand dimension $D_q(r)$ at different value of r. The results are shown in Fig. 2 for $q = -2$ and $+3$. The behaviour is similar for other values of q. We expect the random samples to have $D_q = 3$, the same as the dimension of the ambient space, irrespective of the value of r. We find that $D_q(r)$ has values close to 3, the deviations being less than 7 per cent. At length-scales of $r \sim 30\,h^{-1}$ Mpc, the data and the simulations have $D_q(r)$ values that are quite different from the random samples. We find $D_q \sim 3.2$ and 2.6 for $q = -2$ and $+3$, respectively. The difference from the random samples is larger than the $1 - \sigma$ statistical fluctuations expected for the data. It is clear from this figure that the SDSS data is consistent with the random data at length-scales 70 h^{-1} Mpc and larger. Based on this, we conclude that the transition to homogeneity occurs at a length-scale between 60 and 70 h^{-1} Mpc. We note that the dark matter N-body simulations and the Millennium simulation also exhibit a transition to homogeneity at a similar length-scale as the SDSS galaxies.

We next consider the full spectrum of generalized dimension $D_q(r)$ with $-4 \leq q \leq 4$ at $r = 60$ and 70 h^{-1} Mpc (Fig. 3). At 60 h^{-1} Mpc, the SDSS and random data are consistent for positive q, but they do not agree for $q \leq 0$ where D_q is larger for the SDSS data. This indicates that at 60 h^{-1} Mpc the overdense regions are consistent with homogeneity whereas the underdense regions are not. At 70 h^{-1} Mpc, we find that the SDSS and random data are consistent for the entire q range that we have analysed. This is in keeping with our conclusion that the transition to homogeneity occurs at a length-scale between 60 and 70 h^{-1} Mpc.

Our results are in disagreement with similar analyses by Sylos Labini et al. (2007, 2009a, b) who report long range correlations and persistent fluctuations in the large-scale galaxy distribution, and fail to find a transition to homogeneity. The results of this Letter are in good agreement with an earlier 2D analysis of the SDSS

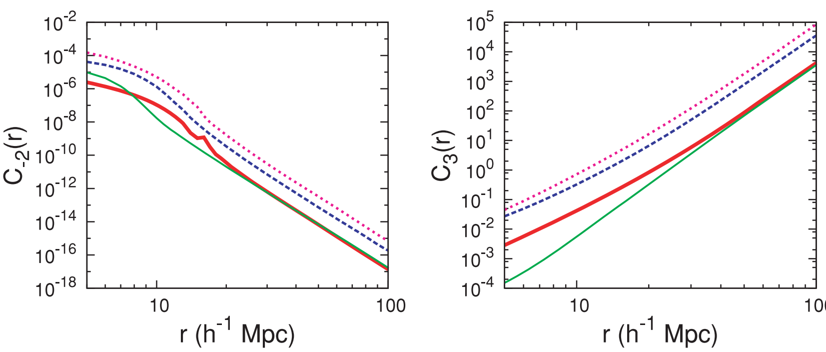

Figure 1. This shows $C_q(r)$ for $q = -2$ (left-hand panel) and $q = 3$ (right-hand panel). The different curves (top to bottom) show the Millennium Run (pink-dotted), dark matter N-body (blue-dashed), SDSS data (red-thick solid) and random data (green-thin solid). The two simulation curves have been scaled arbitrarily for convenience of plotting. The four curves shown here overlap completely at large r if plotted with the same vertical scale. For the N-body and random data, the mean C_q averaged over different independent realizations is shown.

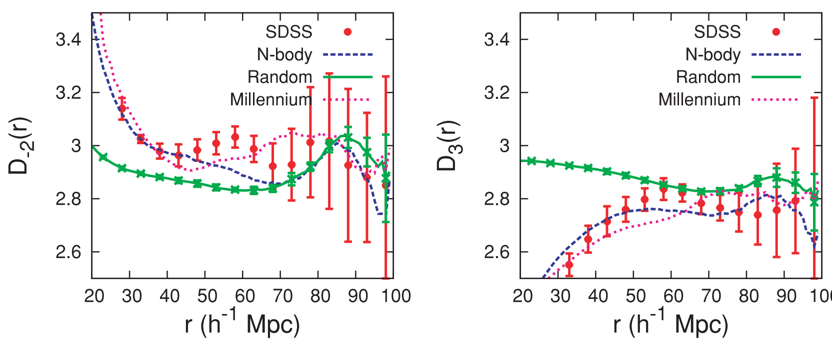

Figure 2. This shows $D_q(r)$ for $q = -2$ (left-hand panel) and $q = 3$ (right-hand panel). The mean and $1 - \sigma$ error bars for the random data and the N-body simulations were, respectively, determined using 10 and 5 independent realizations. We have used the N-body $1 - \sigma$ error bars to estimate the expected statistical fluctuation in the actual SDSS data, and show these on the SDSS results instead of the mean N-body curve. We consider the SDSS data to be consistent with the random data if the random data lies within the $1 - \sigma$ error bars of the SDSS.

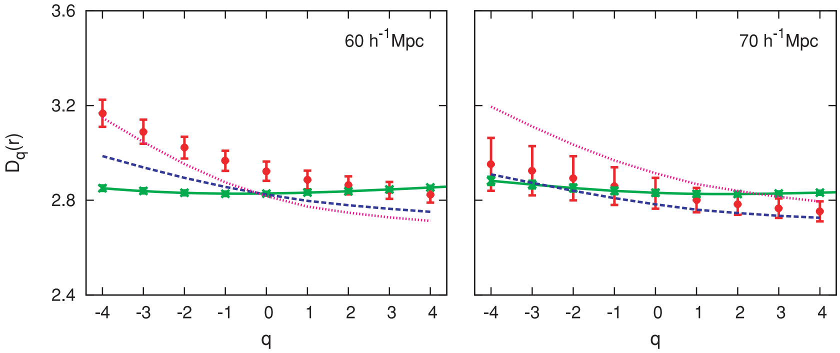

Figure 3. This shows D_q as a function of q for $r = 60$ (left-hand panel) and $70\,h^{-1}$ Mpc (right-hand panel). We show the SDSS data (red points), random data (green solid line), N-body (blue dashed) and Millennium (pink dotted). Five independent N-body realizations were used to estimate the $1-\sigma$ error bars shown on the SDSS data.

main galaxy sample (Yadav et al. 2005) and the SDSS LRG sample (Hogg et al. 2005), both of which find a transition to homogeneity at around $70\,h^{-1}$ Mpc. Note that the LRG sample covers a much larger volume as compared to the main galaxy sample considered in this Letter. A visual inspection of the galaxy distribution reveals that the galaxies appear to be distributed in an interconnected network of filaments encircling voids. This network, referred to as the 'Cosmic Web', appears to fill the entire region covered by galaxy surveys. Bharadwaj, Bhavsar & Sheth (2004) and Pandey & Bharadwaj (2005) have shown that the observed filaments are statistically significant only to length-scales of $70\,h^{-1}$ Mpc and not beyond. Longer filaments, though seen, are produced by chance alignments and are not statistically significant. The good agreement between all of these findings provide strong evidence that the galaxy distribution exhibits a transition to homogeneity at around $70\,h^{-1}$ Mpc.

ACKNOWLEDGMENTS

BP acknowledges Darren Croton for his help in analysing the Millennium catalogue. PS acknowledges financial support of University Grant Commission, India. JY thanks J. S. Bagla, T. R. Seshadri, S. Borgani and F. S. Labini for useful discussions.

The SDSS DR6 data were downloaded from the SDSS skyserver http://cas.sdss.org/dr6/en/. Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho and the Max-Planck Society. The SDSS Web site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck Institute for Astronomy (MPIA), the Max-Planck Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, the United States Naval Observatory and the University of Washington.

The Millennium Run simulation used in this Letter was carried out by the Virgo Supercomputing Consortium at the Computing Centre of the Max-Planck Society in Garching. The semi-analytic galaxy catalogue is publicly available at http://www.mpa-garching.mpg.de/galform/agnpaper.

REFERENCES

Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297
Amendola L., Palladino E., 1999, ApJ, 514, L1
Bagla J. S., Yadav J., Seshadri T. R., 2008, MNRAS, 390, 829
Baryshev Y. V., Bukhmastova Y. L., 2004, Astron. Lett., 30, 444
Best J. S., 2000, ApJ, 541, 519
Bharadwaj S., Gupta A. K., Seshadri T. R., 1999, A&A, 351, 405
Bharadwaj S., Bhavsar S. P., Sheth J. V., 2004, ApJ, 606, 25
Borgani S., 1995, Phys. Rep., 251, 1
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Coleman P. H., Pietronero L., 1992, Phys. Rep., 213, 311
Croton D. J. et al., 2006, MNRAS, 365, 11
Guzzo L., 1997, New Astron., 2, 517
Hatton S., 1999, MNRAS, 310, 1128
Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, ApJ, 624, 54
Jones B. J., Martínez V. J., Saar E., Trimble V., 2005, Rev. Mod. Phys., 76, 1211
Komatsu E. et al., 2009, ApJS, 180, 330
Kurokawa T., Morikawa M., Mouri H., 2001, A&A, 370, 358
Martinez V. J., Jones B. J. T., 1990, MNRAS, 242, 517
Pan J., Coles P., 2000, MNRAS, 318, L51
Pandey B., Bharadwaj S., 2005, MNRAS, 357, 1068
Peebles P. J. E., 1980, The Large Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ, p. 435
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C. The Art of Scientific Computing, 2nd edn. Cambridge Univ. Press, Cambridge
Springel V. et al., 2005, Nat, 435, 629
Stoughton C. et al., 2002, AJ, 123, 485
Strauss M. A. et al., 2002, AJ, 124, 1810
Sylos Labini F., Vasilyev N. L., Baryshev Y. V., 2007, A&A, 465, 23
Sylos Labini F., Vasilyev N. L., Baryshev Y. V., 2009a, Europhys. Lett., 85, 29002
Sylos Labini F., Vasilyev N. L., Pietronero L., Baryshev Y. V., 2009b, Europhys. Lett., 86, 49001
Tikhonov A. V., Makarov D. I., Kopylov A. I., 2000, Bull. Special Astrophys. Obs., 50, 39
Weinberg S., 2008, Cosmology, Oxford University Press, Oxford, UK
Yadav J., Bharadwaj S., Pandey B., Seshadri T. R., 2005, MNRAS, 364, 601

This paper has been typeset from a TEX/LATEX file prepared by the author.